updated: 11/23/00

Hayashi *Econometrics*: Answers to Selected Review Questions

Chapter 2

Section 2.1

1. For n sufficiently large, $|z_n - \alpha| < \varepsilon$, which means $\text{Prob}(|z_n - \alpha| > \varepsilon) = 0$.

2. The equality in the hint implies that $\lim_{n\to\infty} \mathrm{E}[(\mathbf{z}_n - \mathbf{z})'(\mathbf{z}_n - \mathbf{z})] = 0$ if and only if $\lim_{n\to\infty} \mathrm{E}[(z_{nk} - z_k)^2] = 0$ for all k.

Section 2.2

6. Because there is a one-to-one mapping between $(\mathbf{g}_{i-1}, \dots, \mathbf{g}_1)$ and $(\mathbf{z}_{i-1}, \dots, \mathbf{z}_1)$ (i.e., the value of $(\mathbf{g}_{i-1}, \dots, \mathbf{g}_1)$ can be calculated from the value of $(\mathbf{z}_{i-1}, \dots, \mathbf{z}_1)$ and vice versa), $\mathrm{E}(\mathbf{g}_i \mid \mathbf{g}_{i-1}, \dots, \mathbf{g}_1) = \mathrm{E}(\mathbf{g}_i \mid \mathbf{z}_{i-1}, \dots, \mathbf{z}_1) = \mathrm{E}(\mathbf{z}_i - \mathbf{z}_{i-1} \mid \mathbf{z}_{i-1}, \dots, \mathbf{z}_1) = \mathbf{0}$.

7.

$$
\begin{aligned}
&\mathrm{E}(g_i \mid g_{i-1}, \dots, g_2) \\
&= \mathrm{E}[\mathrm{E}(\varepsilon_i \cdot \varepsilon_{i-1} \mid \varepsilon_{i-1}, \dots, \varepsilon_1) \mid g_{i-1}, \dots, g_2] \quad \text{(by Law of Iterated Expectations)} \\
&= \mathrm{E}[\varepsilon_{i-1}\, \mathrm{E}(\varepsilon_i \mid \varepsilon_{i-1}, \dots, \varepsilon_1) \mid g_{i-1}, \dots, g_2] \quad \text{(by linearity of conditional expectations)} \\
&= 0 \quad \text{(since } \{\varepsilon_i\} \text{ is independent white noise).}
\end{aligned}
$$

8. Let $x_i \equiv r_{i1}$. Since $(x_{i-1}, \dots, x_2)$ $(i \geq 3)$ has less information than $(y_{i-2}, \dots, y_1)$, we have

$$\mathrm{E}(x_i \mid x_{i-1}, \dots, x_2) = \mathrm{E}[\mathrm{E}(x_i \mid y_{i-2}, \dots, y_1) \mid x_{i-1}, \dots, x_2] \text{ for } i \geq 3.$$

It is easy to show by the Law of Iterated Expectations that $\mathrm{E}(x_i \mid y_{i-2}, \dots, y_1) = 0$.

Section 2.3

1. We have shown on several occasions that "$\mathrm{E}(\varepsilon_i \mid \mathbf{x}_i) = 0$" is stronger than "$\mathrm{E}(\mathbf{x}_i \cdot \varepsilon_i) = \mathbf{0}$".

2(a) No, $\mathrm{E}(\varepsilon_i^2)$ does not need to exist or to be finite.

3. $\mathbf{S} = \mathrm{E}(\varepsilon_i^2 \mathbf{x}_i \mathbf{x}_i') = \mathrm{E}[\mathrm{E}(\varepsilon_i^2 \mathbf{x}_i \mathbf{x}_i' \mid \mathbf{x}_i)] = \mathrm{E}[\mathrm{E}(\varepsilon_i^2 \mid \mathbf{x}_i)\mathbf{x}_i \mathbf{x}_i']$. The second equality is by Law of Total Expectations. The third equality is by the linearity of conditional expectations.

4. You can use Lemma 2.3(a) to claim, for example, $\text{plim}(\mathbf{b} - \boldsymbol{\beta})'\mathbf{S}_{\mathbf{xx}}(\mathbf{b} - \boldsymbol{\beta}) = 0$ because $(\mathbf{b} - \boldsymbol{\beta})'\mathbf{S}_{\mathbf{xx}}(\mathbf{b} - \boldsymbol{\beta})$ is a continuous function of $\mathbf{b} - \boldsymbol{\beta}$ and $\mathbf{S}_{\mathbf{xx}}$.

5. When you use a consistent estimator $\widehat{\boldsymbol{\beta}}$ to calculate the estimated residual $\widehat{\varepsilon}_i$, (2.3.1) becomes

$$\frac{1}{n}\sum_{i=1}^{n} \widehat{\varepsilon}_{i i}^{2} = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i^2 - 2(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\overline{\mathbf{g}} + (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\mathbf{S}_{\mathbf{xx}}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}).$$

You can use exactly the same argument to claim that the second and the third terms on the RHS converge to zero in probability.

Section 2.4

1. Yes, $SE^*(b_k)$ converges to zero in probability. Consequently, the confidence interval shrinks to a point.

2. By the delta method, $\text{Avar}(\widehat{\lambda}) = (1/\beta)^2 \, \text{Avar}(b)$. The standard error of $\widehat{\lambda}$ is by definition the square root of $1/n$ times the estimated asymptotic variance.

3. Inspection of the formula (2.4.2) for W reveals that the numerical value of W is invariant to $\mathbf{F}$. So, *a fortiori*, the finite-sample distribution and the asymptotic distribution are not affected.

Section 2.5

1. No, because (2.5.1) cannot be calculated by those sample means alone.

2. First, (2.5.4') involves multiplication by n, which is required because it is an asymptotic variance (the variance of the limiting distribution of $\sqrt{n}$ times a sampling error). Second, the middle matrix $\mathbf{B}$ houses estimated errors, rather than error variances.

Section 2.6

5. From the equaion in the hint, we can derive:

$$nR^2 = \frac{1}{\frac{n-K}{n} + \frac{1}{n}(K-1)F}(K-1)F.$$

Since $(K-1)F$ converges in distribution to a random variable, $\frac{1}{n}(K-1)F \to_{\text{p}} 0$ by Lemma 2.4(b). So the factor multiplying $(K-1)F$ on the RHS converges to 1 in probability. Then by Lemma 2.4(c), the asymptotic distribution of the RHS is the same as that of $(K-1)F$, which is chi-squared.

Section 2.8

1. The proof is a routine use of the Law of Total Expectations.

$$\begin{aligned}
&\text{E}(\mathbf{z}_i \cdot \eta_i) \\
&= \text{E}[\text{E}(\mathbf{z}_i \cdot \eta_i \mid \mathbf{x}_i)] \quad \text{(by Law of Total Expectations)} \\
&= \text{E}[\mathbf{z}_i \cdot \text{E}(\eta_i \mid \mathbf{x}_i)] \quad \text{(by linearity of conditional expectations)} \\
&= 0.
\end{aligned}$$

2. The error may be conditionally heteroskedastic, but that doesn't matter asymptotically because all we need from this regression is a consistent estimate of $\boldsymbol{\alpha}$.

Section 2.9

1.

$$\begin{aligned}
&\text{E}[\eta\phi(\mathbf{x})] \\
&= \text{E}\{\text{E}[\eta\phi(\mathbf{x}) \mid \mathbf{x}]\} \quad \text{(by Law of Total Expectations)} \\
&= \text{E}\{\phi(\mathbf{x})\,\text{E}[\eta \mid \mathbf{x}]\} \quad \text{(by linearity of conditional expectations)} \\
&= 0 \quad (\text{since } \text{E}(\eta \mid \mathbf{x}) = \text{E}[y - \text{E}(y \mid \mathbf{x}) \mid \mathbf{x}] = \text{E}(y \mid \mathbf{x}) - \text{E}(y \mid \mathbf{x}) = 0).
\end{aligned}$$

2. Use (2.9.6) to calculate $\mathrm{E}^*(\varepsilon_i \mid \varepsilon_{i-1}, \ldots, \varepsilon_{i-m})$. (I am using E^* for the least squares projection operator.) It is zero. For $\mathrm{E}^*(\varepsilon_i \mid 1, \varepsilon_{i-1}, \ldots, \varepsilon_{i-m})$, use (2.9.7). For white noise processes, $\mu = 0$ and $\boldsymbol{\gamma} = \mathbf{0}$. So $\mathrm{E}^*(\varepsilon_i \mid 1, \varepsilon_{i-1}, \ldots, \varepsilon_{i-m}) = 0$. The conditional expectation, as opposed to the least squares projection, may not be zero. Example 2.4 provides an example.

3. If $\mathrm{E}(y \mid \widetilde{\mathbf{x}}) = \mu + \boldsymbol{\gamma}'\widetilde{\mathbf{x}}$, then y can be written as: $y = \mu + \boldsymbol{\gamma}'\widetilde{\mathbf{x}} + \eta$, $\mathrm{E}(\eta \mid \widetilde{\mathbf{x}}) = 0$. So $\mathrm{Cov}(\widetilde{\mathbf{x}}, y) = \mathrm{Cov}(\widetilde{\mathbf{x}}, \mu + \boldsymbol{\gamma}'\widetilde{\mathbf{x}} + \eta) = \mathrm{Var}(\widetilde{\mathbf{x}})\boldsymbol{\gamma}$. Also, $\mathrm{E}(y) - \boldsymbol{\gamma}'\,\mathrm{E}(\widetilde{\mathbf{x}}) = \mu$. Combine these results with (2.9.7).

4(b) $\boldsymbol{\beta}$.

4(c) The answer is uncertain. For the sake of concreteness, assume $\{y_i, \mathbf{x}_i, \mathbf{z}_i\}$ is i.i.d. Then the asymptotic variance of the estimate of $\boldsymbol{\beta}$ from part (a) is $\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}\,\mathrm{E}(\varepsilon_i^2 \mathbf{x}_i \mathbf{x}_i')\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}$. The asymptotic variance of the estimate of $\boldsymbol{\beta}$ from part (b) is $\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}\,\mathrm{E}[(\mathbf{z}_i'\boldsymbol{\delta} + \varepsilon_i)^2 \mathbf{x}_i \mathbf{x}_i']\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}$. For concreteness, strengthen the orthogonality of $\mathbf{x}_i$ to $(\varepsilon_i, \mathbf{z}_i)$ by the condition that $\mathbf{x}_i$ is independent of $(\varepsilon_i, \mathbf{z}_i)$. Then these two expressions for the asymptotic variance becomes: $\mathrm{E}(\varepsilon_i^2)\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}$ and $\mathrm{E}[(\mathbf{z}_i'\boldsymbol{\delta} + \varepsilon_i)^2]\boldsymbol{\Sigma}_{\mathbf{xx}}^{-1}$. Since ε_i is not necessarily orthogonal to $\mathbf{z}_i$, $\mathrm{E}(\varepsilon_i^2)$ may or may not be greater than $\mathrm{E}[(\mathbf{z}_i'\boldsymbol{\delta} + \varepsilon_i)^2]$.

Section 2.10

1. The last three terms on the RHS of the equation in the hint all converges in probability to μ^2.

2. Let $\mathbf{c}$ be the p-dimensional vector of ones, and

$$\mathbf{d}_n = \left(\frac{n+2}{n-1}, \frac{n+2}{n-2}, \ldots, \frac{n+2}{n-p}\right)'.$$

Then the Box-Pierce /Q can be written as $\mathbf{c}'\mathbf{x}_n$ and the modified Q as $\mathbf{d}_n'\mathbf{x}_n$. Clearly, $\mathbf{a}_n \equiv \mathbf{c} - \mathbf{d}_n$ converges to zero as $n \to \infty$.

Section 2.11

2. You have proved this for general cases in review question 3 of Section 2.9.

3. We can now drop the assumption of constant real rates from the hypothesis of efficient markets. Testing market efficiency then is equivalent to testing whether the inflation forecast error is an m.d.s.

4. If inflation and interest rates are in fractions, then the OLS estimate of the intercept gets divided by 100. The OLS estimate of the interest rate coefficient remains the same. If the inflation rate is in percent per month and the interest rate is in percent per year, then both the intercept and the interest rate coefficient is deflated by a factor of about 12.

5. For the third element of $\mathbf{g}_t$, you can't use the linearity of conditional expectations as in (2.11.7).